Polynomial factor models: non-iterative estimation via method-of-moments

Florian Schuberth[1,2] Rebecca Büchner[3] Karin Schermelleh-Engel[3]
Theo K. Dijkstra[4]

[1]University of Würzburg, Germany

[2]University of Twente, The Netherlands

[3]University of Frankfurt, Germany

[4]Unversity of Groningen, The Netherlands

March 17, 2017

SEM Meeting, Ghent

Overview

1. Polynomial factor model
2. Non-iterative method-of-moments
3. Monte Carlo simulation
4. MoMpoly package

Structural model

Example of a polynomial factor model:

$$
\begin{aligned}
\eta_4 = & \gamma_1\eta_1 + \gamma_2\eta_2 + \gamma_3\eta_3 + \\
& \gamma_{11}(\eta_1^2 - 1) + \gamma_{22}(\eta_2^2 - 1) + \\
& \gamma_{12}(\eta_1\eta_2 - \mathsf{E}(\eta_1\eta_2)) + \gamma_{112}(\eta_1^2\eta_2 - \mathsf{E}(\eta_1^2\eta_2)) \\
& \gamma_{123}(\eta_1\eta_2\eta_3 - \mathsf{E}(\eta_1\eta_2\eta_3)) + \zeta_1.
\end{aligned} \tag{1}
$$

All latent variables (LVs) are standardized, and the structural error term ζ_1 is independent of the LVs on the right-hand side of the equation, i.e., ζ_1 is independent of η_1, η_2, and η_3.

Measurement model

Each LV is connected to at least two indicators and each block of indicators is connected to one LV only:

$$\boldsymbol{y}_i = \boldsymbol{\lambda}_i \eta_i + \boldsymbol{\epsilon}_i. \tag{2}$$

The indicators are standardized and the measurement errors are mutually independent and independent of the η's. The correlation matrix of the indicators of one block can be calculated as

$$\mathrm{E}(\boldsymbol{y}_i \boldsymbol{y}_i') = \boldsymbol{\Sigma}_{ii} = \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' + \boldsymbol{\Theta}_i, \tag{3}$$

where the covariance matrix of the measurement errors $\boldsymbol{\Theta}_i$ is diagonal. The correlation matrix of the indicators of different blocks $(i \neq j)$:

$$\mathrm{E}(\boldsymbol{y}_i \boldsymbol{y}_j') = \boldsymbol{\Sigma}_{ij} = \rho_{ij} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j'. \tag{4}$$

Proxies and weights

To estimate the model parameters, we build a proxy for each latent variable as weighted linear combination of its indicators. The weights used to build the proxy for latent variable i are obtained as

$$\hat{\boldsymbol{w}}_i \propto \sum_{j \neq i} e_{ij} \boldsymbol{S}_{ij} \boldsymbol{w}_j \text{ (one-step weights)}, \tag{5}$$

where $\boldsymbol{w}_j$ is an arbitrary vector of the same length as $\boldsymbol{y}_j$ and $e_{ij} = \text{sign}(\boldsymbol{w}_i' \boldsymbol{S}_{ij} \boldsymbol{w}_j)$. Both weight vectors $\boldsymbol{w}_j$ and $\hat{\boldsymbol{w}}_i$ are scaled: $\boldsymbol{w}_j' \boldsymbol{S}_{jj} \boldsymbol{w}_j = 1$ and $\hat{\boldsymbol{w}}_i' \boldsymbol{S}_{ii} \hat{\boldsymbol{w}}_i = 1$.
The probability limit of $\hat{\boldsymbol{w}}_i$ is

$$\text{plim}(\hat{\boldsymbol{w}}_i) = \bar{\boldsymbol{w}}_i = \boldsymbol{\lambda}_i / \sqrt{\boldsymbol{\lambda}_i' \boldsymbol{\Sigma}_{ii} \boldsymbol{\lambda}_i} = \boldsymbol{\lambda}_i / c_i. \tag{6}$$

Estimation of the loadings

Calculate the correction factor $\hat{c}_i$ such that the squared difference between the off-diagonal elements of

$$\boldsymbol{S}_{ii} \text{ and } \hat{c}_i \hat{\boldsymbol{w}}_i \hat{c}_i \hat{\boldsymbol{w}}_i{}' \tag{7}$$

is minimized. As a result, we obtain

$$\hat{c}_i = \sqrt{\frac{\hat{\boldsymbol{w}}_i'(\boldsymbol{S}_{ii} - \text{diag}(\boldsymbol{S}_{ii}))\hat{\boldsymbol{w}}_i}{\hat{\boldsymbol{w}}_i'(\hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i' - \text{diag}(\hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i'))\hat{\boldsymbol{w}}_i}}, \tag{8}$$

where $\text{diag}(\boldsymbol{S}_{ii})$ is the diagonal matrix of $\boldsymbol{S}_{ii}$.
Since $\text{plim}(\hat{c}_i) = \sqrt{\boldsymbol{\lambda}_i' \boldsymbol{\Sigma}_{ii} \boldsymbol{\lambda}_i}$, the factor loadings can be consistently estimated by $\hat{c}_i \hat{\boldsymbol{w}}_i = \hat{\boldsymbol{\lambda}}_i$.

Relationship between latent variables and proxies

We define a population proxy:

$$\bar{\eta}_i = \bar{\boldsymbol{w}}_i' \boldsymbol{y}_i = (\bar{\boldsymbol{w}}_i' \boldsymbol{\lambda}_i)\eta_i + \bar{\boldsymbol{w}}_i' \boldsymbol{\epsilon}_i = Q_i \eta_i + \delta_i, \tag{9}$$

where Q_i (quality) is the correlation between the proxy and its latent variable. The δ's (as the ϵ's) have zero means and are mutually independent and independent of the η's.
Replacing $\boldsymbol{\lambda}_i$ by $c_i \bar{\boldsymbol{w}}_i$ in Q_i, we obtain

$$Q_i = \bar{\boldsymbol{w}}_i' \boldsymbol{\lambda}_i = c_i \bar{\boldsymbol{w}}_i' \bar{\boldsymbol{w}}_i. \tag{10}$$

We can estimate the quality by

$$\hat{Q}_i = \hat{c}_i \hat{\boldsymbol{w}}_i' \hat{\boldsymbol{w}}_i. \tag{11}$$

Relationship between latent variables and proxies

Relationship between the correlation of the proxies and the correlation between the latent variables:

$$E(\bar{\eta}_i\bar{\eta}_j) = Q_i Q_j E(\eta_i\eta_j), \tag{12}$$

where $E(\bar{\eta}_i\bar{\eta}_j)$ is estimated by the sample covariance between the proxies i and j.

Using PLS weights, this is already known and published:

- for linear models by [Dijkstra, T.K. & Henseler, J., 2015], and
- for polynomial models by [Dijkstra, T. K. & Schermelleh-Engel, K., 2014].

New:

- the use of one-step weights
- implementation in the MoMpoly package in R [Schuberth, F. & Dijkstra, T. K.& Schamberger, T., 2017]

Non-iterative method-of-moments

Starting point is a (regression) equation with a two-way interaction term:

$$\eta_3 = \gamma_1\eta_1 + \gamma_2\eta_2 + \gamma_{12}(\eta_1\eta_2 - \mathsf{E}(\eta_1\eta_2)) + \zeta_1. \tag{13}$$

The γ's can be obtained by solving the following moment equations:

$$\begin{pmatrix} \mathsf{E}(\eta_1\eta_3) \\ \mathsf{E}(\eta_2\eta_3) \\ \mathsf{E}(\eta_1\eta_2\eta_3) \end{pmatrix} = \begin{pmatrix} 1 & \mathsf{E}(\eta_1\eta_2) & \mathsf{E}(\eta_1^2\eta_2) \\ & 1 & \mathsf{E}(\eta_1\eta_2^2) \\ & & \mathsf{E}(\eta_1^2\eta_2^2) - \mathsf{E}(\eta_1\eta_2)^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_{12} \end{pmatrix}, \tag{14}$$

where the moments are given by:

$$\mathsf{E}(\bar{\eta}_i\bar{\eta}_j) = Q_iQ_j\mathsf{E}(\eta_i\eta_j), \tag{15}$$

$$\mathsf{E}(\bar{\eta}_i^2\bar{\eta}_j) = Q_i^2Q_j\mathsf{E}(\eta_i^2\eta_j), \tag{16}$$

$$\mathsf{E}(\bar{\eta}_i^2\bar{\eta}_j^2) = Q_i^2Q_j^2(\mathsf{E}(\eta_i^2\eta_j^2) - 1) + 1, \text{ and} \tag{17}$$

$$\mathsf{E}(\bar{\eta}_i\bar{\eta}_j\bar{\eta}_k) = Q_iQ_jQ_k\mathsf{E}(\eta_i\eta_j\eta_k). \tag{18}$$

Non-iterative method-of-moments

So far, **no distributional assumptions are necessary**. We only assume that

- the moments exist,
- the measurement errors are mutually independent and independent of the η's, and
- that the structural error term is independent of the η's of the right-hand side of the equation.

Non-iterative method-of-moments

Adding quadratic terms to the equation with the two-way interaction term:

$$\begin{aligned}\eta_3 =& \gamma_1\eta_1 + \gamma_2\eta_2 + \\ & \gamma_{11}(\eta_1^2 - 1) + \gamma_{12}(\eta_1\eta_2 - \mathrm{E}(\eta_1\eta_2)) + \gamma_{22}(\eta_2^2 - 1) + \zeta_1. \qquad (19)\end{aligned}$$

Higher moments are required and therefore more assumptions are necessary, in particular, assumptions about the higher-order moments of the δ's:

$$\mathrm{E}(\bar{\eta}_i{}^3) = Q_i^3\,\mathrm{E}(\eta_i^3) + \mathrm{E}(\delta_i^3), \qquad (20)$$

$$\mathrm{E}(\bar{\eta}_i^4) = Q_i^4\,\mathrm{E}(\eta_i^4) + 6Q_i^2(1 - Q_i^2) + \mathrm{E}(\delta_i^4), \text{ and} \qquad (21)$$

$$\mathrm{E}(\bar{\eta}_i^3\bar{\eta}_j) = Q_i^3 Q_j\,\mathrm{E}(\eta_i^3\eta_j) + 3\,\mathrm{E}(\bar{\eta}_i\bar{\eta}_j)(1 - Q_i^2). \qquad (22)$$

Assumptions about the moments of the δ's

In the following, we assume that δ_i, has the same higher-order moments as the normal distribution, i.e.,

$$\begin{aligned} \mathrm{E}(\delta_i^3) =& 0, \text{ and} && (23) \\ \mathrm{E}(\delta_i^4) =& 3\left[\mathrm{var}(\delta_i)\right]^2 = 3(1 - Q_i^2)^2. && (24) \end{aligned}$$

Joint normality of all exogenous variables

Instead of using the *general* moments, we can assume joint normality for all exogenous variables ϵ's, ζ, and exogenous η's . Considering again the equation of example above:

$$\begin{pmatrix} \mathsf{E}(\eta_1\eta_3) \\ \mathsf{E}(\eta_2\eta_3) \\ \mathsf{E}(\eta_1\eta_2\eta_3) \end{pmatrix} = \begin{pmatrix} 1 & \mathsf{E}(\eta_1\eta_2) & \mathsf{E}(\eta_1^2\eta_2) \\ & 1 & \mathsf{E}(\eta_1\eta_2^2) \\ & & \mathsf{E}(\eta_1^2\eta_2^2) - \mathsf{E}(\eta_1\eta_2)^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_{12} \end{pmatrix}. \tag{25}$$

Assuming joint normality leads to:

$$\begin{pmatrix} \mathsf{E}(\eta_1\eta_3) \\ \mathsf{E}(\eta_2\eta_3) \\ \mathsf{E}(\eta_1\eta_2\eta_3) \end{pmatrix} = \begin{pmatrix} 1 & \rho_{12} & 0 \\ & 1 & 0 \\ & & 1 + \rho_{12}^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_{12} \end{pmatrix}, \tag{26}$$

where $\rho_{12} = \mathsf{E}(\bar{\eta}_1\bar{\eta}_2)/(Q_1 Q_2)$. Covariances between the LVs of the RHS and the dependent LV are calculated as above. In principle every recursive model can be estimated in this way.

First simulation

Setting:

- Structural model:

$$\eta_3 = 0.3\eta_1 + 0.4\eta_2 + \\ 0.12(\eta_1^2 - 1) + 0.15(\eta_1\eta_2 - 0.3) + 0.1(\eta_2^2 - 1) + \zeta_1. \qquad (27)$$

- Correlation between η_1 and η_2 is set to $\rho_{12} = 0.3$
- Each latent variable is connected with 3 indicators with $\boldsymbol{\lambda}_i' = \begin{pmatrix} 0.9 & 0.85 & 0.8 \end{pmatrix}$
- Sample size of $N = 400$ and 500 runs
- Estimators: MoMpoly with one-step weights, and Latent Moderated Structural Equations (LMS) [Klein, A. & Moosbrugger, H., 2000]

Results

Para.	true	mean MoMpoly[1]	mean LMS[2]	sd MoMpoly[1]	sd LMS[2]
γ_1	0.30	0.298	0.297	0.048	0.047
γ_2	0.40	0.403	0.402	0.047	0.045
γ_{11}	0.12	0.118	0.117	0.044	0.041
γ_{12}	0.15	0.147	0.147	0.062	0.059
γ_{22}	0.10	0.104	0.101	0.042	0.039
ρ_{12}	0.30	0.300	0.299	0.052	0.052
λ_{11}	0.90	0.898	0.899	0.044	0.017
λ_{12}	0.85	0.854	0.852	0.052	0.018
λ_{13}	0.80	0.795	0.800	0.058	0.022

[1]Inadmissible results are removed, therefore the results are based on 484 estimations.
[2]No inadmissible results were produced.

Second simulation

Our approach can also deal with several recursive equations. For demonstration, we run a second simulation using the following model:

$$\eta_3 = 0.5\eta_1 - 0.4\eta_2 + \zeta_1 \tag{28}$$

$$\eta_4 = 0.35\eta_1 + 0.2\eta_3 + 0.1(\eta_3^2 - 1) + \zeta_2 \tag{29}$$

$$\eta_5 = 0.3\eta_2 - 0.25\eta_4 + 0.3(\eta_3\eta_4 - 0.333) + \zeta_3 \tag{30}$$

- $\mathrm{E}(\eta_1\eta_2) = 0.3$
- Each latent variable is connected with 3 indicators with $\boldsymbol{\lambda}_i' = \begin{pmatrix} 0.9 & 0.85 & 0.8 \end{pmatrix}$
- Sample size of $N = 400$ and 500 runs
- Estimator: MoMpoly with one-step weights

Results

Para.	true	mean[1]	sd[1]
γ_{31}	0.500	0.502	0.049
γ_{32}	-0.400	-0.395	0.053
γ_{41}	0.350	0.347	0.056
γ_{43}	0.200	0.201	0.058
γ_{433}	-0.100	-0.099	0.038
γ_{52}	0.300	0.301	0.050
γ_{54}	-0.250	-0.247	0.052
γ_{534}	0.300	0.298	0.052
ρ_{12}	0.300	0.299	0.051
λ_{51}	0.900	0.895	0.054
λ_{51}	0.850	0.846	0.061
λ_{51}	0.800	0.803	0.068

[1] Inadmissible results are removed, therefore the results are based on 475 estimations.

The MoMpoly package

Model specification in *lavaan* syntax [Rosseel, Y., 2012], e.g.,

```
model='
    ETA3~ETA1+ETA2+ETA1.ETA1+ETA1.ETA2+ETA2.ETA2
    ETA1 =~ y1+y2+y3
    ETA2 =~ y4+y5+y6
    ETA3 =~ y7+y8+y9
'
```

The model can be estimated by the `MoMpoly` function from the *MoMpoly* package:

```
MoMpoly(model, data)
```

Capabilities of the MoMpoly package

It is capable to deal with:

- single equations using
  - the *general* moments (however, skewness and kurtosis of the δ's from the normal distribution are used)
  - the moments from the multivariate normal distribution (all exogenous variables: ϵ's, ζ, and exogenous η's are jointly normally distributed)
- correlated measurement error within a block
- several recursive equations using the moments from the multivariate normal distribution (all exogenous variables: ϵ's, and ζ's, and exogenous η's are jointly normally distributed) where the parameter estimates
  - are retrieved from reduced form coefficient matrix
  - are obtained from replacing equation by equation
- several recursive equations using the *general* moments. The parameters are estimated equation by equation (limited to LVs up to higher-order 3)
- single-indicator latent variables

Further options in the MoMpoly function and functions

```
MoMpoly(model, data,criterionload=criterion.geometric,
        criterionpath=criterion.diff,normality=F,
        weightFun=weightFun.onestep,
        KindOfEstimation=approach.none,...)
```

Further options in the MoMpoly function:

- different ways to obtain the correction factor $\hat{c}_i$
- different correction factor for loadings and path coefficients
- different weights: one-step weights, PLS weights, or prescribed weights

Further functions:

- `MoMpoly.boot()`: combination of the `MoMpoly` function with the *boot* package
- `MoMpoly.jack()`: function to obtain Jackknife standard errors

Further research

- Elaboration of the measurement model via instrumental variable techniques
- Other higher-order moments for delta than those from the normal distribution)
- Robustness against model misspecifications

Thank you!
Questions/Comments?

References

Dijkstra, T. K. & Schermelleh-Engel, K., (2014)
Consistent partial least squares for nonlinear structural equation models
Psychometrika 79(4) 585 – 604.

Dijkstra, T.K. & Henseler, J. (2015)
Consistent and asymptotically normal PLS estimators for linear structural equations
Computational Statistics & Data Analysis, 81, 10 – 23.

Klein, A. & Moosbrugger, H. (2000)
Maximum likelihood estimation of latent interaction effects with the LMS method
Psychometrika, 65(4), 457 – 474.

Rosseel, Y. (2012)
lavaan: An R Package for Structural Equation Modeling
Journal of Statistical Software, 48(2), 1 – 36.

Schuberth, F., Schamberger, T. & Dijkstra, T. K. (2017)
MoMpoly package (version 0.1.3.) *available upon request*.

Results first simulation

Para.	true	mean	sd
γ_1	0.300	0.298	0.048
γ_{11}	0.120	0.118	0.044
γ_{12}	0.150	0.147	0.062
γ_2	0.400	0.403	0.047
γ_{22}	0.100	0.104	0.042
ρ_{12}	0.300	0.300	0.052
λ_{31}	0.900	0.896	0.034
λ_{32}	0.850	0.850	0.039
λ_{33}	0.800	0.797	0.041
λ_{11}	0.900	0.898	0.044
λ_{12}	0.850	0.854	0.052
λ_{13}	0.800	0.795	0.058
λ_{21}	0.900	0.898	0.037
λ_{22}	0.850	0.850	0.044
λ_{23}	0.800	0.800	0.049

Results second simulation

Para.	true	mean	sd
γ_{31}	0.500	0.502	0.049
γ_{32}	-0.400	-0.395	0.053
γ_{41}	0.350	0.347	0.056
γ_{43}	0.200	0.201	0.058
γ_{433}	-0.100	-0.099	0.038
γ_{52}	0.300	0.301	0.050
γ_{534}	0.300	0.298	0.052
γ_{54}	-0.250	-0.247	0.052
ρ_{12}	0.300	0.299	0.051
λ_{31}	0.900	0.904	0.036
λ_{32}	0.850	0.847	0.040
λ_{33}	0.800	0.797	0.045

Para	true	mean	sd
λ_{41}	0.900	0.898	0.040
λ_{42}	0.850	0.850	0.043
λ_{43}	0.800	0.800	0.048
λ_{51}	0.900	0.895	0.054
λ_{52}	0.850	0.846	0.061
λ_{53}	0.800	0.803	0.068
λ_{11}	0.900	0.897	0.035
λ_{12}	0.850	0.851	0.040
λ_{13}	0.800	0.799	0.042
λ_{21}	0.900	0.897	0.041
λ_{22}	0.850	0.848	0.046
λ_{23}	0.800	0.803	0.049